

## **RAW SEQUENCE LISTING**

**Loaded by SCORE, no errors detected.**

Application Serial Number: 10599851

Source: OPAP

Date Processed by SCORE: 8/13/09

***ENTERED***

<110> APPLICANT: UAB Research Foundation  
KAPPES, John C.  
MULKY, Alok  
WU, Xiaoyun  
<120> TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR IDENTIFYING  
COMPOUNDS THAT INHIBIT HIV-1 SUBUNIT-SPECIFIC REVERSE  
TRANSCRIPTASE  
<130> FILE REFERENCE: 21085.0123P1  
  
<140> CURRENT APPLICATION NUMBER: 10599851  
<141> CURRENT FILING DATE: 0001-01-01  
<150> PRIOR APPLICATION NUMBER: PCT/US05/18335  
<151> PRIOR FILING DATE: 2005-06-02  
<150> PRIOR APPLICATION NUMBER: 60/573,918  
<151> PRIOR FILING DATE: 2004-05-24  
<150> PRIOR APPLICATION NUMBER: 60/668,858  
<151> PRIOR FILING DATE: 2005-04-06  
<160> NUMBER OF SEQ ID NOS: 21  
<170> SOFTWARE: FastSEQ for Windows Version 4.0  
  
<210> SEQ ID NO 1  
<211> LENGTH: 858  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence; note =  
synthetic construct  
<400> SEQUENCE: 1  
60 gtttaaacgc caccatggag caggcccccg aggaccaggg ccccccagagg gagccccaca  
120 acgagtggac cctggagctg ctggaggagc tgaagaggga ggccgtgagg cacttccca  
180 ggccctggct gcacggcctg gcccagcaca tctacgagac ctacggcgac acctggcccg  
240 gcgtggaggg catcatcagg atcctgcagc agctgctgtt catccacttc aggatcggct  
300 gccagcacag caggatcggc atcatccagc agaggaggc caggaggaac ggcgccagca  
360 ggagctagtt taaacactgc acagagagac aggctaattt tttagggaaa atttggcctt  
420 ccaacaaagg gaggccaggg aattttctcc agaacaggcc agagccaaaca gccccaccccg  
480 cagagagcct cgggttcgga gaggagatag cccctcccc gaaacaagag ccgaaggaaa  
540 aggagttata ccccttaacc tccctcaaat cactcttgg cagcgacccc tagtcacagt  
600 aagaataggg ggacagctaa tagaaggccct gtttagacaca ggagcagatg atacagtgtt  
660 agaagatata aatttaccag gaaaatggaa accaaaaatg atagggggaa ttgggtggct  
720 tatcaaagta agacagtatg atcaaatact tatagaaatt tgtggaaaaa aggctatagg  
780 gacagtatta gtaggaccta cacctatcaa cataattggg agaaatatgt tgactcagat  
840 tggttgtact taaaatttc caattagtcc tattgaaact gtaccagtaa aattaaagcc  
858 aggaatggat ggtccaaa  
  
<210> SEQ ID NO 2  
<211> LENGTH: 96

<212> TYPE: PRT  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence; note = synthetic construct  
<400> SEQUENCE: 2  
Met Glu Gln Ala Pro Glu Asp Gln Gly Pro Pro Arg Glu Pro Tyr Asn  
1 5 10 15  
Ala Trp Thr Leu Glu Leu Leu Glu Leu Lys Ser Glu Ala Val Arg  
20 25 30  
His Phe Pro Arg Val Trp Leu His Gly Leu Gly Gln His Ile Tyr Glu  
35 40 45  
Thr Tyr Gly Asp Thr Trp Ala Gly Val Glu Ala Ile Ile Arg Ile Leu  
50 55 60  
Gln Gln Leu Leu Phe Ile His Phe Arg Ile Gly Cys Gln His Ser Arg  
65 70 75 80  
Ile Gly Ile Thr Arg Gln Arg Arg Ala Arg Asn Gly Ala Ser Arg Ser  
85 90 95

<210> SEQ ID NO 3  
<211> LENGTH: 315  
<212> TYPE: DNA  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence; note = synthetic construct  
<400> SEQUENCE: 3  
gtttaaacgc caccatggag caggcccccg aggaccaggg ccccccagagg gagccccaca  
60  
acgagtggac cctggagctg ctggaggagc tgaagaggga ggccgtgagg cacttccccca  
120  
ggccctggct gcacggcctg ggccagcaca tctacgagac ctacggcgac acctggggccg  
180  
gcgtggaggc catcatcagg atcctgcagc agctgctgtt catccacttc aggatcggct  
240  
gccagcacag caggatcggc atcatccagc agaggaggc caggaggaac ggcgccagca  
300  
ggagctagtt taaac  
315

<210> SEQ ID NO 4  
<211> LENGTH: 440  
<212> TYPE: PRT  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence; note = synthetic construct  
<400> SEQUENCE: 4  
Pro Ile Ser Pro Ile Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met  
1 5 10 15  
Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys  
20 25 30  
Ala Leu Val Glu Ile Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser  
35 40 45  
Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys  
50 55 60  
Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu  
65 70 75 80  
Asn Lys Arg Thr Gln Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His  
85 90 95  
Pro Ala Gly Leu Lys Lys Lys Ser Val Thr Val Leu Asp Val Gly

	100	105	110												
Asp	Ala	Tyr	Phe	Ser	Val	Pro	Leu	Asp	Glu	Asp	Phe	Arg	Lys	Tyr	Thr
	115		120									125			
Ala	Phe	Thr	Ile	Pro	Ser	Ile	Asn	Asn	Glu	Thr	Pro	Gly	Ile	Arg	Tyr
	130		135									140			
Gln	Tyr	Asn	Val	Leu	Pro	Gln	Gly	Trp	Lys	Gly	Ser	Pro	Ala	Ile	Phe
	145		150						155				160		
Gln	Ser	Ser	Met	Thr	Lys	Ile	Leu	Glu	Pro	Phe	Arg	Lys	Gln	Asn	Pro
	165		170						170				175		
Asp	Ile	Val	Ile	Tyr	Gln	Tyr	Met	Asp	Asp	Leu	Tyr	Val	Gly	Ser	Asp
	180		185									190			
Leu	Glu	Ile	Gly	Gln	His	Arg	Thr	Lys	Ile	Glu	Glu	Leu	Arg	Gln	His
	195		200						200			205			
Leu	Leu	Arg	Trp	Gly	Leu	Thr	Thr	Pro	Asp	Lys	Lys	His	Gln	Lys	Glu
	210		215									220			
Pro	Pro	Phe	Leu	Trp	Met	Gly	Tyr	Glu	Leu	His	Pro	Asp	Lys	Trp	Thr
	225		230						235			240			
Val	Gln	Pro	Ile	Val	Leu	Pro	Glu	Lys	Asp	Ser	Trp	Thr	Val	Asn	Asp
	245		250						250			255			
Ile	Gln	Lys	Leu	Val	Gly	Lys	Leu	Asn	Trp	Ala	Ser	Gln	Ile	Tyr	Pro
	260		265						265			270			
Gly	Ile	Lys	Val	Arg	Gln	Leu	Cys	Lys	Leu	Leu	Arg	Gly	Thr	Lys	Ala
	275		280						280			285			
Leu	Thr	Glu	Val	Ile	Pro	Leu	Thr	Glu	Glu	Ala	Glu	Leu	Glu	Leu	Ala
	290		295						295			300			
Glu	Asn	Arg	Glu	Ile	Leu	Lys	Glu	Pro	Val	His	Gly	Val	Tyr	Tyr	Asp
	305		310						310			315			320
Pro	Ser	Lys	Asp	Leu	Ile	Ala	Glu	Ile	Gln	Lys	Gln	Gly	Gln	Gly	Gln
	325		330						330			335			
Trp	Thr	Tyr	Gln	Ile	Tyr	Gln	Glu	Pro	Phe	Lys	Asn	Leu	Lys	Thr	Gly
	340		345						345			350			
Lys	Tyr	Ala	Arg	Met	Arg	Gly	Ala	His	Thr	Asn	Asp	Val	Lys	Gln	Leu
	355		360						360			365			
Thr	Glu	Ala	Val	Gln	Lys	Ile	Thr	Thr	Glu	Ser	Ile	Val	Ile	Trp	Gly
	370		375						375			380			
Lys	Thr	Pro	Lys	Phe	Lys	Leu	Pro	Ile	Gln	Lys	Glu	Thr	Trp	Glu	Thr
	385		390						390			395			400
Trp	Trp	Thr	Glu	Tyr	Trp	Gln	Ala	Thr	Trp	Ile	Pro	Glu	Trp	Glu	Phe
	405		410						410			415			
Val	Asn	Thr	Pro	Pro	Leu	Val	Lys	Leu	Trp	Tyr	Gln	Leu	Glu	Lys	Glu
	420		425						425			430			
Pro	Ile	Val	Gly	Ala	Glu	Thr	Phe								
	435		440												

<210> SEQ\_ID NO 5  
<211> LENGTH: 440  
<212> TYPE: PRT  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence; note = synthetic construct  
<400> SEQUENCE: 5  
Pro Ile Ser Pro Ile Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met  
1 5 10 15  
Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys  
20 25 30  
Ala Leu Val Glu Ile Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser  
35 40 45  
Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys  
50 55 60  
Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu

65	70	75	80
Asn	Lys	Arg	Thr
Gln	Asp	Phe	Trp
Glu	Val	Gln	Leu
Leu	Gly	Ile	Pro
85	90	95	
Pro	Ala	Gly	Leu
Lys	Lys	Lys	Ser
Val	Thr	Val	Leu
Asp	Val	Gly	
100	105	110	
Asp	Ala	Tyr	Phe
Ser	Val	Pro	Leu
Asp	Glu	Asp	Phe
Arg	Lys	Tyr	Thr
115	120	125	
Ala	Phe	Thr	Ile
Pro	Ser	Ile	Asn
Asn	Glu	Thr	Pro
Gly	Ile	Arg	Tyr
130	135	140	
Gln	Tyr	Asn	Val
Val	Leu	Pro	Gln
Gly	Trp	Lys	Gly
Ser	Pro	Ala	Ile
145	150	155	160
Gln	Ser	Ser	Met
Thr	Lys	Ile	Leu
Glu	Pro	Phe	Arg
Lys	Gln	Ile	Arg
165	170	175	
Asp	Ile	Val	Ile
Tyr	Gln	Tyr	Met
Asp	Asp	Leu	Tyr
Val	Gly	Gly	Ser
180	185	190	
Leu	Glu	Ile	Gly
Gln	His	Arg	Thr
195	200	205	
Leu	Leu	Arg	Trp
Gly	Leu	Thr	Thr
Pro	Asp	Lys	Lys
His	Gln	His	Gln
210	215	220	
Pro	Pro	Phe	Leu
Trp	Met	Gly	Tyr
Glu	Leu	His	Pro
225	230	235	240
Val	Gln	Pro	Ile
Ile	Val	Leu	Pro
Glu	Lys	Asp	Ser
245	250	255	
Ile	Gln	Lys	Leu
Val	Gly	Lys	Leu
Asn	Trp	Ala	Ser
260	265	270	
Gly	Ile	Lys	Val
275	280	285	
Leu	Thr	Glu	Val
Ile	Pro	Leu	Thr
Glu	Glu	Ala	Glu
290	295	300	
Glu	Asn	Arg	Glu
Ile	Leu	Lys	Glu
305	310	315	320
Pro	Ser	Lys	Asp
Leu	Ile	Ala	Glu
Ile	Gln	Lys	Gln
325	330	335	
Trp	Thr	Tyr	Gln
Ile	Tyr	Gln	Glu
340	345	350	
Lys	Tyr	Ala	Arg
355	360	365	
Met	Arg	Gly	Ala
Arg	Gly	Ala	His
370	375	380	
Thr	Glu	Ala	Val
Gln	Lys	Ile	Thr
385	390	395	400
Trp	Trp	Thr	Glu
Tyr	Trp	Gln	Ala
405	410	415	
Val	Asn	Thr	Pro
Pro	Leu	Val	Lys
420	425	430	
Pro	Ile	Val	Gly
435	440		

<210> SEQ\_ID NO 6  
<211> LENGTH: 170  
<212> TYPE: PRT  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence; note =  
synthetic construct  
<400> SEQUENCE: 6  
Lys Glu Gly His Gln Met Lys Glu Cys Thr Glu Arg Gln Ala Asn Phe  
1 5 10 15  
Leu Gly Lys Ile Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe Leu  
20 25 30  
Gln Ser Arg Pro Glu Pro Thr Ala Pro Pro Glu Glu Ser Phe Arg Cys

35	40	45													
Gly	Glu	Glu	Lys	Thr	Thr	Pro	Pro	Gln	Lys	Pro	Glu	Gln	Thr	Asp	Lys
50						55				60					
Glu	Leu	Tyr	Pro	Leu	Ala	Ser	Leu	Arg	Ser	Leu	Phe	Gly	Gln	Arg	Pro
65						70				75				80	
Leu	Val	Thr	Ile	Lys	Ile	Gly	Gly	Gln	Leu	Lys	Glu	Ala	Leu	Leu	Asp
						85				90				95	
Thr	Gly	Ala	Asp	Asp	Thr	Val	Leu	Glu	Asp	Met	Ser	Leu	Pro	Gly	Lys
						100				105				110	
Trp	Lys	Pro	Lys	Met	Ile	Gly	Gly	Ile	Gly	Gly	Phe	Ile	Lys	Val	Arg
						115				120				125	
Gln	Tyr	Asp	Gln	Ile	Pro	Ile	Glu	Ile	Cys	Gly	His	Lys	Ala	Ile	Gly
						130				135				140	
Thr	Val	Leu	Ile	Gly	Pro	Thr	Pro	Val	Asn	Ile	Ile	Gly	Arg	Asn	Leu
						145				150				155	
Leu	Thr	Gln	Ile	Gly	Cys	Thr	Leu	Asn	Phe						160
						165									170

<210> SEQ ID NO 7  
 <211> LENGTH: 511  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence; note = synthetic construct  
 <400> SEQUENCE: 7  
 aaaggaagga caccaaatga aagaatgcac tgagagacag gctaatttt tagggaaaat  
 60 ctggccttcc cacaaggaa gcccaggaa ctttctccag agcagaccag agccaacagc  
 120 cccaccagaa gagagcttca ggtgtgggaa ggagaaaaca actccccctc agaagccgga  
 180 gcagacagac aaggaactgt atccttttagc ttccctcaga tcactctttg gcaacgaccc  
 240 ctcgtcacaa taaagatagg ggggcagcta aaggaagctc tattagatac aggagcagat  
 300 gatacagtat tagaagacat gagtttgcca ggaaaatgga agccaaaaat gataggggaa  
 360 attggaggtt ttatcaaagt aagacagtat gatcagatac ctatagaaat ctgtgggcat  
 420 aaagctatag gtacagtatt aataggacca acacctgtca acataattgg aagaaatctg  
 480 ttgacacaga ttgggtgcac tttaaatttt c  
 511

<210> SEQ ID NO 8  
 <211> LENGTH: 4  
 <212> TYPE: PRT  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Description of Artificial Sequence; note = synthetic construct  
 <400> SEQUENCE: 8  
 Tyr Met Asp Asp  
 1

<210> SEQ ID NO 9  
 <211> LENGTH: 4  
 <212> TYPE: PRT  
 <213> ORGANISM: Artificial Sequence  
 <220> FEATURE: